

## SEQUENCE LISTING

<110> Wicher, Kryzysztof B.  
 Holst, Olof Peder  
 Hachem, Maher Youssef Abou  
 Karlsson, Eva Margareta Nordberg  
 Hreggvidsson, Gudmundur O.

<120> Thermostable Cellulase

<130> P5099PC00

<150> PCT/IS01/00012

<151> 2001-06-15

<150> 09/594,884

<151> 2000-06-15

<160> 10 ✓

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1877

<212> DNA

<213> Rhodothermus marinus

<220>

<221> CDS

<222> (710) .. (1495)

<223>

<400> 1

```

gggcgagtcg aacgtcgggc ggcctcgca gatcgaacgg ctcttccac gctgattcgc      60
tgcgtgcgca cgcactgggc gctgggcgag tttcccttct acttcgtcca gattgccctt      120
tacgactacg ggccgaacgc ccgcacgcag cgtgtccggg aagcgcagtt gcgcaccatg      180
tgcgcgtgcc tcatacgggt atggtcgtga ccaccgatgt gggcgacgac cacaacacta      240
cctccggcgc gacaacgtga agtggcgagc gcctggcctc tgggcgctgg ccaacgacta      300
cggcttccgg gatgtggtct actccgggtc catctacgaa cgcattggaac gtgaggatgg      360
ccgcctgcgg ctctatttcc gctatgccga agggggactg gtccctgcgc cggccgacgg      420
cggcggcgag ttcgtgattg cgggaccgga ccgtgtcttc caccggcgc gggtaggggt      480
cgaaggcgaa acgctggtgg tgtggagtcc gcgcgtctcc gatccgcagg cggcgcgcta      540
tggctggtcc aatacgccgc acgccacgct gttcaaccgg gccggcctgc ccgcttcgcc      600
cttccgcacc gacgactggc cggaaggcga ctgagcgacg caaccggtgc ttgcatgcga      660
caggggcact tcgtaccttg aagtgccctt tgtcatttca atggaataa atg aac gtc      718
Met Asn Val
1

```

atg	cgt	gcg	gta	ctg	gtc	ctg	agc	ctg	tta	ttg	ctg	ttt	gga	tgc	gac	766
Met	Arg	Ala	Val	Leu	Val	Leu	Ser	Leu	Leu	Leu	Leu	Phe	Gly	Cys	Asp	
5						10					15					
tg	g	c	t	a	t	t	a	a	a	a	a	a	a	a	a	814
Trp	Leu	Phe	Pro	Asp	Gly	Asp	Asn	Gly	Lys	Glu	Pro	Glu	Pro	Glu	Pro	
20					25					30					35	
gag	ccg	acc	gtc	gag	ctg	tgc	gga	cgc	tgg	gac	gcg	cgc	gat	gtg	gcc	862
Glu	Pro	Thr	Val	Glu	Leu	Cys	Gly	Arg	Trp	Asp	Ala	Arg	Asp	Val	Ala	
			40						45					50		
ggg	ggg	cgc	tac	cgg	gtg	atc	aac	aac	gta	tgg	ggc	gcg	gag	acc	gcc	910
Gly	Gly	Arg	Tyr	Arg	Val	Ile	Asn	Asn	Val	Trp	Gly	Ala	Glu	Thr	Ala	
			55					60					65			
cag	tgc	att	gag	gtc	gga	ctg	gaa	acg	ggc	aac	ttc	acg	atc	aca	cgg	958
Gln	Cys	Ile	Glu	Val	Gly	Leu	Glu	Thr	Gly	Asn	Phe	Thr	Ile	Thr	Arg	
		70					75					80				
gcc	gat	cac	gac	aac	ggc	aac	aac	gtg	gcc	gcc	tat	ccg	gcc	atc	tac	1006
Ala	Asp	His	Asp	Asn	Gly	Asn	Asn	Val	Ala	Ala	Tyr	Pro	Ala	Ile	Tyr	
	85					90					95					
ttc	ggg	tgc	cac	tgg	ggc	gcc	tgc	acg	agc	aac	tcg	gga	ttg	ccg	cgg	1054
Phe	Gly	Cys	His	Trp	Gly	Ala	Cys	Thr	Ser	Asn	Ser	Gly	Leu	Pro	Arg	
100					105					110					115	
cgc	gtg	cag	gag	ctg	tcc	gac	gtg	cgc	acg	agc	tgg	acg	ctc	acg	ccg	1102
Arg	Val	Gln	Glu	Leu	Ser	Asp	Val	Arg	Thr	Ser	Trp	Thr	Leu	Thr	Pro	
				120					125					130		
atc	acg	acg	ggc	cgc	tgg	aac	gcc	gcc	tac	gac	atc	tgg	ttc	agt	ccc	1150
Ile	Thr	Thr	Gly	Arg	Trp	Asn	Ala	Ala	Tyr	Asp	Ile	Trp	Phe	Ser	Pro	
			135				140						145			
gtc	acg	aac	tcc	ggc	aac	ggc	tac	agc	ggc	ggc	gcc	gag	ctg	atg	atc	1198
Val	Thr	Asn	Ser	Gly	Asn	Gly	Tyr	Ser	Gly	Gly	Ala	Glu	Leu	Met	Ile	
		150					155					160				
tgg	ctg	aac	tgg	aac	ggc	ggc	gtg	atg	ccg	ggc	ggc	agc	cgc	gtg	gcc	1246
Trp	Leu	Asn	Trp	Asn	Gly	Gly	Val	Met	Pro	Gly	Gly	Ser	Arg	Val	Ala	
	165				170						175					
acc	gtg	gaa	ctg	gcc	ggg	gcc	acc	tgg	gaa	gtc	tgg	tat	gcc	gac	tgg	1294
Thr	Val	Glu	Leu	Ala	Gly	Ala	Thr	Trp	Glu	Val	Trp	Tyr	Ala	Asp	Trp	
180					185					190					195	
gac	tgg	aac	tac	atc	gcc	tac	cgg	cgc	acg	acg	ccc	acc	acg	tcg	gtg	1342
Asp	Trp	Asn	Tyr	Ile	Ala	Tyr	Arg	Arg	Thr	Thr	Pro	Thr	Thr	Ser	Val	
				200					205					210		
agc	gag	ctg	gac	ctg	aag	gcc	ttc	atc	gac	gac	gcg	gtc	gcc	cgc	ggc	1390
Ser	Glu	Leu	Asp	Leu	Lys	Ala	Phe	Ile	Asp	Asp	Ala	Val	Ala	Arg	Gly	
			215				220						225			
tac	atc	cgg	ccg	gag	tgg	tat	ctg	cat	gcg	gtg	gag	acg	ggc	ttc	gaa	1438
Tyr	Ile	Arg	Pro	Glu	Trp	Tyr	Leu	His	Ala	Val	Glu	Thr	Gly	Phe	Glu	
		230					235					240				

ctc tgg gag ggc ggg gcc ggt ctg cga agc gcc gat ttt tcc gta acg 1486  
 Leu Trp Glu Gly Gly Ala Gly Leu Arg Ser Ala Asp Phe Ser Val Thr  
           245                                  250                                  255

gtg cag tag cctgtcacac gggcaccagc gtaggccaga gaagcacccg 1535  
 Val Gln  
 260

tcggggcgct tatgcgggcg gccgtcgctt tgtgcctgac tttgtagtgc gctacggagg 1595

cgtcagccgg cgtgggtgcgt ttccctggag gcgcttcgctt tcgtgccgga cgagaactta 1655

tccgaaaagc aaggagagaga cctggtaagc cagggctttg cactgcgcac cagaagcacg 1715

ggttttagt aagccacgaa ggcgtccgcc gaagtggcct caccgcgcac ggcacttcgc 1775

tgcgctccgt gcctgactac gagcgcgttg gccgtgcttc aaacggcagg ctataaagct 1835

atacaggtcg aaatcgcgaa cttctccgac acttacaac ac 1877

<210> 2  
 <211> 261  
 <212> PRT  
 <213> Rhodothermus marinus

<400> 2

Met Asn Val Met Arg Ala Val Leu Val Leu Ser Leu Leu Leu Leu Phe  
 1                                  5                                  10                                  15

Gly Cys Asp Trp Leu Phe Pro Asp Gly Asp Asn Gly Lys Glu Pro Glu  
                                   20                                  25                                  30

Pro Glu Pro Glu Pro Thr Val Glu Leu Cys Gly Arg Trp Asp Ala Arg  
                                   35                                  40                                  45

Asp Val Ala Gly Gly Arg Tyr Arg Val Ile Asn Asn Val Trp Gly Ala  
                                   50                                  55                                  60

Glu Thr Ala Gln Cys Ile Glu Val Gly Leu Glu Thr Gly Asn Phe Thr  
 65                                  70                                  75                                  80

Ile Thr Arg Ala Asp His Asp Asn Gly Asn Asn Val Ala Ala Tyr Pro  
                                   85                                  90                                  95

Ala Ile Tyr Phe Gly Cys His Trp Gly Ala Cys Thr Ser Asn Ser Gly  
                                   100                                  105                                  110

Leu Pro Arg Arg Val Gln Glu Leu Ser Asp Val Arg Thr Ser Trp Thr  
                                   115                                  120                                  125

Leu Thr Pro Ile Thr Thr Gly Arg Trp Asn Ala Ala Tyr Asp Ile Trp  
 130 135 140

Phe Ser Pro Val Thr Asn Ser Gly Asn Gly Tyr Ser Gly Gly Ala Glu  
 145 150 155 160

Leu Met Ile Trp Leu Asn Trp Asn Gly Gly Val Met Pro Gly Gly Ser  
 165 170 175

Arg Val Ala Thr Val Glu Leu Ala Gly Ala Thr Trp Glu Val Trp Tyr  
 180 185 190

Ala Asp Trp Asp Trp Asn Tyr Ile Ala Tyr Arg Arg Thr Thr Pro Thr  
 195 200 205

Thr Ser Val Ser Glu Leu Asp Leu Lys Ala Phe Ile Asp Asp Ala Val  
 210 215 220

Ala Arg Gly Tyr Ile Arg Pro Glu Trp Tyr Leu His Ala Val Glu Thr  
 225 230 235 240

Gly Phe Glu Leu Trp Glu Gly Gly Ala Gly Leu Arg Ser Ala Asp Phe  
 245 250 255

Ser Val Thr Val Gln  
 260

<210> 3  
 <211> 786  
 <212> DNA  
 <213> Rhodothermus marinus

<400> 3  
 atgaacgtca tgcgtgcggt actggtcctg agcctgttat tgctgtttgg atgcgactgg 60  
 ctctttcccg atggcgacaa cggaaaggaa ccggagcctg agcccgagcc gaccgtcgag 120  
 ctgtgcggac gctgggacgc gcgcgatgtg gccggggggc gctaccgggt gatcaacaac 180  
 gtatggggcg cggagaccgc ccagtgcatt gaggtcggac tggaaacggg caacttcacg 240  
 atcacacggg ccgatcacga caacggcaac aacgtggccg cctatccggc catctacttc 300  
 gggtgccact ggggcgcctg cacgagcaat tcgggattgc cgcggcgcgt gcaggagctg 360  
 tccgacgtgc gcacgagctg gacgctcacg ccgatcacga cgggccgctg gaatgccgcc 420  
 tacgacatct gggtcagtcc cgtcacgaat tccggcaacg gctacagcgg cggcgccgag 480  
 ctgatgatct ggctgaactg gaacggcggc gtgatgccgg gcggcagccg cgtggccacc 540

gtggaactgg ccggggccac ctgggaagtc tggtatgccg actgggactg gaattacatc 600  
gcctaccggc gcacgacgcc caccacgtcg gtgagcgagc tggacctgaa ggccttcac 660  
gacgacgcgg tcgcccgcgg ctacatccgg ccggagtggg atctgcatgc ggtggagacg 720  
ggcttcgaac tctgggaggg cggggccggg ctgcgaagcg ccgatttttc cgtaacggtg 780  
cagtag 786

<210> 4  
<211> 1134  
<212> DNA  
<213> *Pyrococcus furiosus*

<220>  
<221> exon  
<222> (82) .. (1041)  
<223>

<400> 4  
gaatgtcggg gcagattacg aagagagaag agctcttttt tataccttat attcaaatct 60  
aacaaatttg gaggatgtcc a atg agc aag aaa aag ttc gtc atc gta tct 111  
Met Ser Lys Lys Lys Phe Val Ile Val Ser  
1 5 10  
atc tta aca atc ctt tta gta cag gca ata tat ttt gta gaa aag tat 159  
Ile Leu Thr Ile Leu Leu Val Gln Ala Ile Tyr Phe Val Glu Lys Tyr  
15 20 25  
cat acc tct gag gac aag tca act tca aat acc tca tct aca cca ccc 207  
His Thr Ser Glu Asp Lys Ser Thr Ser Asn Thr Ser Ser Thr Pro Pro  
30 35 40  
caa aca aca ctt tcc act acc aag gtt ctc aag att aga tac cct gat 255  
Gln Thr Thr Leu Ser Thr Thr Lys Val Leu Lys Ile Arg Tyr Pro Asp  
45 50 55  
gac ggt gag tgg cca gga gct cct att gat aag gat ggt gat ggg aac 303  
Asp Gly Glu Trp Pro Gly Ala Pro Ile Asp Lys Asp Gly Asp Gly Asn  
60 65 70  
cca gaa ttc tac att gaa ata aac cta tgg aac att ctt aat gct act 351  
Pro Glu Phe Tyr Ile Glu Ile Asn Leu Trp Asn Ile Leu Asn Ala Thr  
75 80 85 90  
gga ttt gct gag atg acg tac aat tta acc agc ggc gtc ctt cac tac 399  
Gly Phe Ala Glu Met Thr Tyr Asn Leu Thr Ser Gly Val Leu His Tyr  
95 100 105  
gtc caa caa ctt gac aac att gtc ttg agg gat aga agt aat tgg gtg 447  
Val Gln Gln Leu Asp Asn Ile Val Leu Arg Asp Arg Ser Asn Trp Val  
110 115 120  
cat gga tac ccc gaa ata ttc tat gga aac aag cca tgg aat gca aac 495  
His Gly Tyr Pro Glu Ile Phe Tyr Gly Asn Lys Pro Trp Asn Ala Asn  
125 130 135

tac gca act gat ggc cca ata cca tta ccc agt aaa gtt tca aac cta	543
Tyr Ala Thr Asp Gly Pro Ile Pro Leu Pro Ser Lys Val Ser Asn Leu	
140 145 150	
aca gac ttc tat cta aca atc tcc tat aaa ctt gag ccc aag aac gga	591
Thr Asp Phe Tyr Leu Thr Ile Ser Tyr Lys Leu Glu Pro Lys Asn Gly	
155 160 165 170	
ctg cca att aac ttc gca ata gaa tcc tgg tta acg aga gaa gct tgg	639
Leu Pro Ile Asn Phe Ala Ile Glu Ser Trp Leu Thr Arg Glu Ala Trp	
175 180 185	
aga aca aca gga att aac agc gat gag caa gaa gta atg ata tgg att	687
Arg Thr Thr Gly Ile Asn Ser Asp Glu Gln Glu Val Met Ile Trp Ile	
190 195 200	
tac tat gac gga tta caa ccg gct ggc tcc aaa gtt aag gag att gta	735
Tyr Tyr Asp Gly Leu Gln Pro Ala Gly Ser Lys Val Lys Glu Ile Val	
205 210 215	
gtc cca ata ata gtt aac gga aca cca gta aat gct aca ttt gaa gta	783
Val Pro Ile Ile Val Asn Gly Thr Pro Val Asn Ala Thr Phe Glu Val	
220 225 230	
tgg aag gca aac att ggt tgg gag tat gtt gca ttt aga ata aag acc	831
Trp Lys Ala Asn Ile Gly Trp Glu Tyr Val Ala Phe Arg Ile Lys Thr	
235 240 245 250	
cca atc aaa gag gga aca gtg aca att cca tac gga gca ttt ata agt	879
Pro Ile Lys Glu Gly Thr Val Thr Ile Pro Tyr Gly Ala Phe Ile Ser	
255 260 265	
gtt gca gcc aac att tca agc tta cca aat tac aca gaa ctt tac tta	927
Val Ala Ala Asn Ile Ser Ser Leu Pro Asn Tyr Thr Glu Leu Tyr Leu	
270 275 280	
gag gac gtg gag att gga act gag ttt gga acg cca agc act acc tcc	975
Glu Asp Val Glu Ile Gly Thr Glu Phe Gly Thr Pro Ser Thr Thr Ser	
285 290 295	
gcc cac cta gag tgg tgg atc aca aac ata aca cta act cct cta gat	1023
Ala His Leu Glu Trp Trp Ile Thr Asn Ile Thr Leu Thr Pro Leu Asp	
300 305 310	
aga cct ctt att tcc taa atttcggcaa cctgggaatt atcaagttaa	1071
Arg Pro Leu Ile Ser	
315	
agaaaagggtg gagttgctaa agaattcaaa gaaaatttga aaagtaactt ttattgtgat	1131
ctc	1134

<210> 5  
 <211> 319  
 <212> PRT  
 <213> *Pyrococcus furiosus*

<220>

<221> SIGNAL  
 <222> (1)..(19)  
 <223>

<220>  
 <221> DOMAIN  
 <222> (28)..(49)  
 <223> Linker Moiety

<220>  
 <221> DOMAIN  
 <222> (50)..(319)  
 <223> Catalytic Domain

<400> 5

Met Ser Lys Lys Lys Phe Val Ile Val Ser Ile Leu Thr Ile Leu Leu  
 1 5 10 15

Val Gln Ala Ile Tyr Phe Val Glu Lys Tyr His Thr Ser Glu Asp Lys  
 20 25 30

Ser Thr Ser Asn Thr Ser Ser Thr Pro Pro Gln Thr Thr Leu Ser Thr  
 35 40 45

Thr Lys Val Leu Lys Ile Arg Tyr Pro Asp Asp Gly Glu Trp Pro Gly  
 50 55 60

Ala Pro Ile Asp Lys Asp Gly Asp Gly Asn Pro Glu Phe Tyr Ile Glu  
 65 70 75 80

Ile Asn Leu Trp Asn Ile Leu Asn Ala Thr Gly Phe Ala Glu Met Thr  
 85 90 95

Tyr Asn Leu Thr Ser Gly Val Leu His Tyr Val Gln Gln Leu Asp Asn  
 100 105 110

Ile Val Leu Arg Asp Arg Ser Asn Trp Val His Gly Tyr Pro Glu Ile  
 115 120 125

Phe Tyr Gly Asn Lys Pro Trp Asn Ala Asn Tyr Ala Thr Asp Gly Pro  
 130 135 140

Ile Pro Leu Pro Ser Lys Val Ser Asn Leu Thr Asp Phe Tyr Leu Thr  
 145 150 155 160

Ile Ser Tyr Lys Leu Glu Pro Lys Asn Gly Leu Pro Ile Asn Phe Ala  
 165 170 175

Ile Glu Ser Trp Leu Thr Arg Glu Ala Trp Arg Thr Thr Gly Ile Asn  
 180 185 190

Ser Asp Glu Gln Glu Val Met Ile Trp Ile Tyr Tyr Asp Gly Leu Gln  
 195 200 205

Pro Ala Gly Ser Lys Val Lys Glu Ile Val Val Pro Ile Ile Val Asn  
 210 215 220

Gly Thr Pro Val Asn Ala Thr Phe Glu Val Trp Lys Ala Asn Ile Gly  
 225 230 235 240

Trp Glu Tyr Val Ala Phe Arg Ile Lys Thr Pro Ile Lys Glu Gly Thr  
 245 250 255

Val Thr Ile Pro Tyr Gly Ala Phe Ile Ser Val Ala Ala Asn Ile Ser  
 260 265 270

Ser Leu Pro Asn Tyr Thr Glu Leu Tyr Leu Glu Asp Val Glu Ile Gly  
 275 280 285

Thr Glu Phe Gly Thr Pro Ser Thr Thr Ser Ala His Leu Glu Trp Trp  
 290 295 300

Ile Thr Asn Ile Thr Leu Thr Pro Leu Asp Arg Pro Leu Ile Ser  
 305 310 315

<210> 6  
 <211> 12  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> HSV-Tag

<400> 6

Ser Gln Pro Glu Leu Ala Pro Glu Asp Pro Glu Asp  
 1 5 10

<210> 7  
 <211> 29  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Oligonucleotide

<400> 7



aggactccat atgaccgtcg agctgtcgg

29

<210> 8

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 8

acctgagaag cttctgcacc gttacgga

28

<210> 9

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 9

agctccatat gtgcgactgg ctctttcc

28

<210> 10

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Oligonucleotide

<400> 10

acctgagaag cttctgcacc gttacgga

28